THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Fusing simulated GEDI, ICESat-2 and NISAR data for regional aboveground biomass mapping

Citation for published version:
Silva, CA, Duncanson, L, Hancock, S, Neuenschwander, A, Thomas, N, Hofton, M, Fatoyinbo, L, Simard, M, Marshak, CZ, Armston, J, Lutchke, S & Dubayah, R 2021, 'Fusing simulated GEDI, ICESat-2 and NISAR data for regional aboveground biomass mapping', *Remote Sensing of Environment*, vol. 253, pp. 112234. https://doi.org/10.1016/j.rse.2020.112234

Digital Object Identifier (DOI):
10.1016/j.rse.2020.112234

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Remote Sensing of Environment

Publisher Rights Statement:
© 2020 The Author(s). Published by Elsevier Inc.

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

OPEN ACCESS

Download date: 05. May. 2023

Remote Sensing of Environment 253 (2021) 112234

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Fusing simulated GEDI, ICESat-2 and NISAR data for regional aboveground biomass mapping

Carlos Alberto Silva [a,b,*], Laura Duncanson [a], Steven Hancock [d], Amy Neuenschwander [e], Nathan Thomas [c,f], Michelle Hofton [a], Lola Fatoyinbo [c], Marc Simard [g], Charles Z. Marshak [g], John Armston [a], Scott Lutchke [c], Ralph Dubayah [a]

[a] Department of Geographical Sciences, University of Maryland, College Park, MD 20740, USA
[b] School of Forest Resources and Conservation, University of Florida, Gainesville, FL 3261, USA
[c] Biosciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20707, USA
[d] School of GeoSciences, University of Edinburgh, United Kingdom of Great Britain and Northern Ireland
[e] Applied Research Laboratories, University of Texas at Austin, Austin, TX 78712, USA
[f] Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD 20740, USA
[g] NASA-Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA

ARTICLE INFO

Keywords:
Biomass
Lidar
Mapping
Fusion
Temperate forest
L-band SAR

ABSTRACT

Accurate mapping of forest aboveground biomass (AGB) is critical for better understanding the role of forests in the global carbon cycle. NASA's current GEDI and ICESat-2 missions as well as the upcoming NISAR mission will collect synergistic data with different coverage and sensitivity to AGB. In this study, we present a multi-sensor data fusion approach leveraging the strength of each mission to produce wall-to-wall AGB maps that are more accurate and spatially comprehensive than what is achievable with any one sensor alone. Specifically, we calibrate a regional L-band radar AGB model using the sparse, simulated spaceborne lidar AGB estimates. We assess our data fusion framework using simulations of GEDI, ICESat-2 and NISAR data from airborne laser scanning (ALS) and UAVSAR data acquired over the temperate high AGB forest and complex terrain in Sonoma County, California, USA. For ICESat-2 and GEDI missions, we simulate two years of data coverage and AGB at footprint level are estimated using realistic AGB models. We compare the performance of our fusion framework when different combinations of the sparse simulated GEDI and ICESat-2 AGB estimates are used to calibrate our regional L-band AGB models. In addition, we test our framework at Sonoma using (a) 1-ha square grid cells and (b) similarly sized irregularly shaped objects. We demonstrate that the estimated mean AGB across Sonoma is more accurately estimated using our fusion framework than using GEDI or ICESat-2 mission data alone, either with a regular grid or with irregular segments as mapping units. This research highlights methodological opportunities for fusing new and upcoming active remote sensing data streams toward improved AGB mapping through data fusion.

1. Introduction

Forest ecosystems cover approximately one third of the Earth's land and play a major role in the global carbon budget (FAO, 2018). Accurate measurements of forest aboveground biomass (AGB) over large spatial scales are crucial not only to improve our understanding of the global carbon cycle and achieve effective carbon emission mitigation strategies (Chen et al., 2016; Hese et al., 2005; Houghton et al., 2009), but also for many other societal and scientific tasks such as sustainable forest management and monitoring forest ecosystem productivity and conservation (Hudak et al., 2009; Tian et al., 2012; Silva et al., 2016). However, it is not practical nor cost effective to use field studies for large regional AGB estimates (Hummel et al., 2011). As such, it is necessary to utilize the range of available remotely sensed data to estimate AGB for carbon science and policy.

Several NASA's missions are collecting data sensitive to forest AGB, with different mission concepts and measurements designs. NASA's GEDI (Global Ecosystem Dynamics Investigation Lidar) is a spaceborne

* Corresponding author at: Department of Geographical Sciences, University of Maryland, College Park, MD 20740, USA.
E-mail address: carlos_engflorestal@outlook.com (C.A. Silva).

https://doi.org/10.1016/j.rse.2020.112234
Received 11 May 2020; Received in revised form 30 September 2020; Accepted 30 November 2020
Available online 10 December 2020
0034-4257/© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

lidar system specifically designed to measure forest vertical structure. GEDI began collecting science data in April 2019 for a nominal two-year mission onboard the International Space Station (ISS). GEDI uses a ~25 m diameter footprint full-waveform lidar data to characterize vegetation structure required to estimate AGB between 51.6 degrees North and South (Dubayah et al., 2020; Duncanson et al., 2020). NASA's ICESat-2 (Ice, Cloud, and Land Elevation Satellite 2) is a free-flying satellite mission for estimating vegetation characteristics and land topography but is mostly focused on ice sheet elevation and sea ice thickness (Neuenschwander and Pitts, 2019). ICESat-2 was launched on September 15th, 2018 and for a planned minimum 3 years mission, with the potential for more years of operation. The Advanced Topographic Laser Altimeter System (ATLAS), instrument on ICESat-2 collects global photon counting lidar data that are used to measure forest structure (e.g. height) and can potentially be used for forest AGB monitoring (Montesano et al., 2015; Neuenschwander and Pitts, 2019; Narine et al., 2019a; Narine et al., 2019b).

In addition to GEDI and ICESat-2, the NASA Indian Space Research Organization (ISRO) Synthetic Aperture Radar (NISAR) mission is a joint project to co-develop and launch a dual frequency Synthetic Aperture Radar (SAR) satellite, which will also provide data sensitive to forest AGB (NISAR, 2020). The NISAR mission is NASA's first orbiting L-band SAR mission. NISAR is expected to be launched in 2022 and will collect wall-to-wall L- band data (and S- band in some regions) globally repeating its orbit every 12 days during its three-year mission, providing high-resolution data capable of mapping and monitoring forest AGB and disturbance. Other spaceborne missions with goals to measure AGB (e.g. European Space Agency (ESA)'s BIOMASS (Carreiras et al., 2017; Quegan et al., 2019) and The German Aerospace Center (DLR)'s TanDEM-L (Krieger et al., 2010; Moreira et al., 2018)) are not discussed in this paper.

GEDI, ICESat-2 and NISAR missions will collect data sensitive to vegetation structure with different strengths and limitations. For instance, GEDI has the highest canopy penetration but does not collect data over the boreal region, whereas ICESat-2 collects data globally with less capability to penetrate dense canopies. In addition to being sampling missions, both GEDI and ICESat-2 use active optical lidar technology (near-IR for GEDI and visible for ICESat-2), and thus are sensitive to atmospheric conditions, such as clouds, as well as solar illumination effects. In contrast, NISAR provides global wall-to-wall coverage even in areas with cloud cover and works equally well during the day and night. The biomass requirements for the NISAR mission are only for areas with AGB $<$ 100 Mg/ha with 1-ha resolution, given L-band limitations at higher biomass regimes (Duncanson et al., 2020). These missions provide potentially complementary datasets and studies focused on algorithm development toward optimal use of these mission datasets for fusion is of great interest.

There are a variety of approaches used for estimating forest AGB from L-band backscatter, and according to Santoro and Cartus (2018), these approaches can be grouped into three main categories: i) parametric empirical regression models; ii) parametric semi-empirical and physically-based models; and, iii) non-parametric models. Moreover, calibrating regional L-band AGB models using the AGB estimates from airborne lidar has been anticipated for almost a decade to improve the calibration, accuracy and range of SAR estimates of AGB (Sun et al., 2011; Zolkos et al., 2013). The global-scale coverage of GEDI and ICESat-2 make such regional calibration applicable across much of Earth's forests. Regional calibration of L-band AGB models mitigates (a) the phenological and environmental effects associated with L-band AGB retrieval associated with fluctuations in the dielectric constant of vegetation (Bouvet et al., 2018) and (b) AGB model saturation, especially when the model parameters are determined with field measurements outside the region of study (Bouvet et al., 2018; Zolkos et al., 2013). Such regional calibration of forest models with GEDI in particular have been investigated with respect to TanDEM-X PolInSAR canopy height (Qi et al., 2019a; Qi et al., 2019b) and Landsat biomass models (Saarela et al., 2018, Deo et al., 2017).

In forest areas, GEDI and ICESat-2 data may be aggregated using either grid or object segments for AGB modeling from L-band backscatter data. Although several studies have used remote sensing data and grid based approaches for modeling and wall-to-wall forest structure mapping at a variety of scales, recent studies have shown improvements in AGB estimation accuracies when an object oriented approach is used (Mareya et al., 2018; Silveira et al., 2019; Zhang et al., 2018). For instance, Silveira et al. (2019) compared the accuracy of grid and object-based approaches in the estimation of AGB using a Random Forest algorithm (Liaw and Wiener, 2015) in the Brazilian Atlantic forest, populated with environmental, remote sensing and terrain variables. In their study, they found a notable percent increase of ~53% in terms of R^2 and ~40% in terms of absolute RMSE when compared to the traditional grid-based approach. While the object-based approach might improve AGB estimation accuracies, trade-offs still exist between estimation accuracies and data processing complexity. For instance, it is unclear how much the grid or object-based approach would impact AGB estimation accuracies when combining GEDI, ICESat-2 and NISAR data. Therefore, the development and test of multi-sensor data fusion approaches specifically for the use of GEDI, ICESat-2 and NISAR mission datasets will help inform the production of AGB maps that may have higher accuracy, more complete spatial coverage, and high spatial resolution than maps from any one mission alone.

The overall goal of this paper is to develop and test multi-sensor data fusion approaches for improving forest AGB estimation and wall-to-wall mapping in Sonoma County, California, USA. We applied simulated GEDI and ICESat-2 mission-like AGB estimates through different fusion scenarios for AGB wall-to-wall mapping. We focus on two fusion approaches, i) a 1-ha regular grid-based approach and ii) an object-oriented approach using NISAR image segmentation. In both approaches, GEDI and ICESat-2 estimates of AGB are used to train a NISAR L-band backscatter AGB model. In the grid-based approach, GEDI and ICESat-2 AGB estimates are used to calibrate simulated NISAR L-band backscatter data to produce a wall-to-wall 1-ha AGB map. In the object approach, the GEDI and ICESat-2 data are identical to the grid-based approach but development and application of lidar to the backscatter model is done at an object-level derived from NISAR L-band image segmentation. For each of these approaches we explore three scenarios for training L-band SAR with lidar AGB estimates to NISAR L-band backscatter: 1) using GEDI AGB estimates, 2) using ICESat-2 AGB estimates, and 3) using both GEDI and ICESat-2. We also tested the impact of ICESat-2 photon return rates on suitability for ICESat-2 as an input for fusion AGB mapping.

2. Material and method

2.1. Study area

The study area of this paper is in Sonoma County, California, United States (Fig. 1a–b). Sonoma County's climate is characterized as Mediterranean, and the study area contains three major climate zones: Marine, Coastal Cool, and Coastal Warm (Sonoma County Climatic Zones, 2007). The annual precipitation on average ranges from approximately 500 mm in the southeastern county to 700 to 1000 mm in central and northern valley areas, while the overall average monthly temperature ranges between 7.3 °C to 22.6 °C over a year (Berman, 2006). The topography is varied, including coastal terraces, distinctive valleys, and several mountain ranges (Berman, 2006). Elevation ranges from 0 m to 1366 m, and the predominant soil types in the study area are Alfisols, Ultisols and Inceptisols (Godfrey, 2015). The vegetation of the study area consists primarily of conifers, deciduous and mixed forests, wetlands, herb and shrub with a large gradient of forest structure and AGB (Fig. 1b and d).

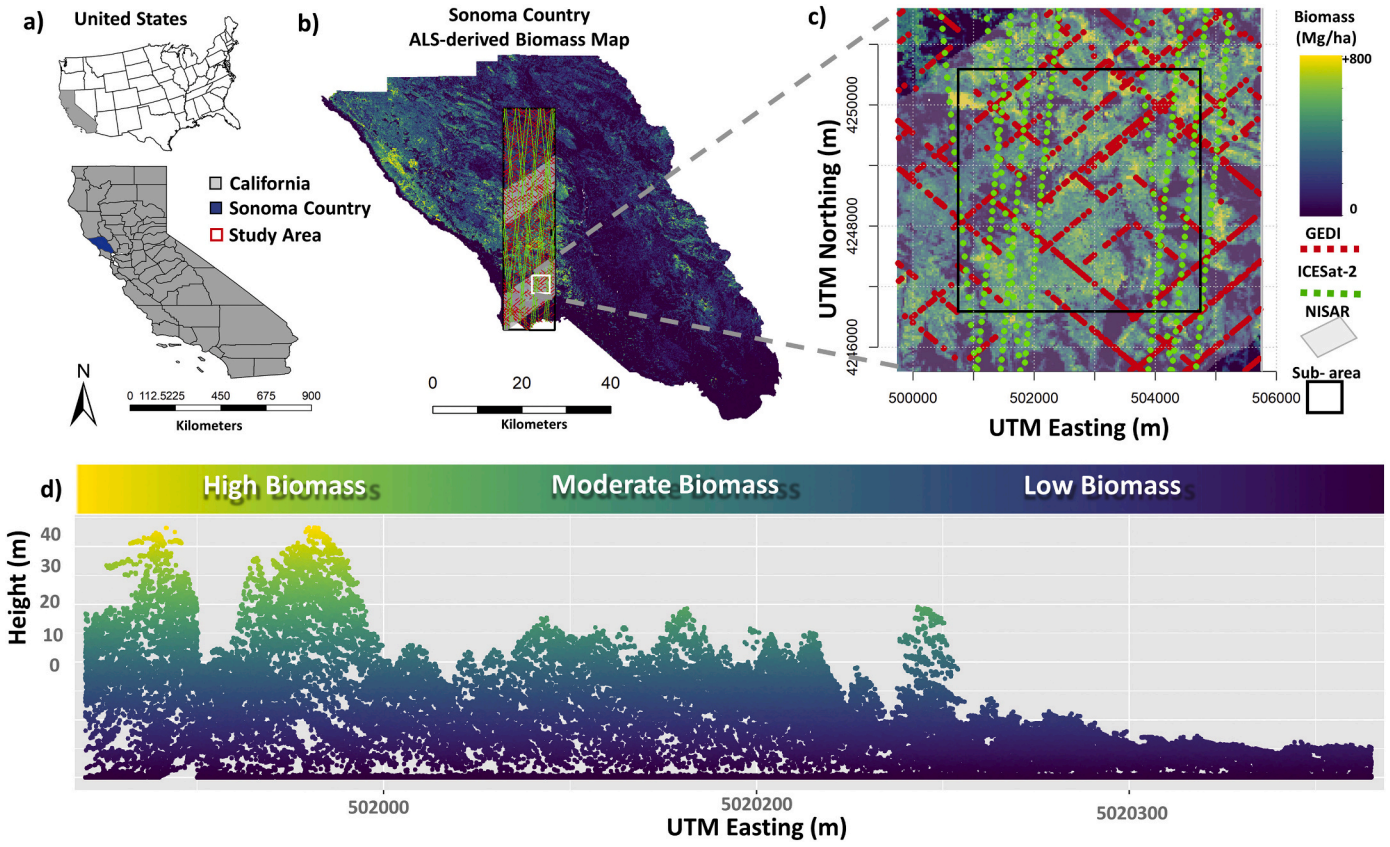

Fig. 1. Study area; a) United States, State of California and Sonoma Country boundaries; b) Airborne lidar (ALS)-derived AGB; c) Illustration of simulated GEDI, ICESat-2 and NISAR data; and d) Sample of ALS-derived vertical profile (100 m × 90 m) with gradients of forest AGB from c).

2.2. Remote sensing data collection and processing

2.2.1. Airborne laser scanning (ALS)

ALS was acquired over Sonoma County between September and November of 2013 using a Leica ALS70 mounted on a Cessna Grand Caravan. The survey covered the whole county (~440,000 ha) at an average point density of ~14 points m^{-2} (Dubayah et al., 2013). A reference AGB map was made by developing a parametric model relating field estimates of AGB from 194 variable radius plots to a suite of ALS derived height and cover metrics. This model was applied to make a wall-to-wall AGB map of Sonoma County at a 30 m spatial resolution (Fig. 1b) (Duncanson et al., 2020b). The calibrated parametric model presented R^2 of ~0.72 and RMSE of ~129 Mg/ha (56.5%). We process the ALS data to simulate GEDI and ICESat-2 across the County; this allows both a richer sample of data representing expected data densities at the end of two years of on-orbit coverage, removes any geolocation uncertainty from on-orbit data, and removes any temporal changes in vegetation structure e.g. from substantial fire activity that occurred between the ALS acquisition and the launch of GEDI and ICESat-2.

2.2.2. Simulated AGB estimates from Global Ecosystem Dynamics Investigation Lidar (GEDI)

NASA's GEDI is a spaceborne lidar system operating from the Japanese Experimental Module's Exposed Facility on the International Space Station. It contains three lasers, emitting 1064 nm pulses, at 242 Hz, that illuminate the Earth's surface with footprints of ~25 m diameter. Two of the lasers are full power, and one is split into two beams, producing a total of four beams, which are optically dithered to produce eight ground tracks; four full-power and four coverage beam tracks, with footprints separated by ~60 m along-track and 600 m across track (Dubayah et al., 2020).

A simulator was developed for GEDI's pre-launch AGB calibration (Hancock et al., 2019). In combination with an orbital track simulator, two years of simulated GEDI waveforms were produced over the study area, including realistic noise and assuming 50% cloud cover. A random stratified sample of 50,000 footprints, covering all range of canopy cover, ground slope and tree height, were chosen from the GEDI simulations, which are independent of the simulated tracks used in this study. Ground detections with a range of processing parameters (smoothing widths, noise threshold, minimum feature widths and with and without noise-tracking) were made across all 50,000 footprints. The smoothing width (applied after calculating noise statistics but before denoising) varied from sigma = 0 cm to sigma = 1 m in steps of 25 cm. The minimum feature width varied from 1 to 7 ranging bins in steps of 2. The noise threshold was set as the mean noise level plus between 2 and 4.5 standard deviations, in steps of 0.5 standard deviations. The ground elevations estimated from the GEDI simulations were compared to those from ALS data and the single combination of parameters with the lowest average RMSE across all footprints across all sites was selected. For these simulations, the denoising parameters were smoothing width = 0.5 m; minimum feature width = 1 bins; noise threshold = mean + 2.0*std. and noise-tracking = yes, which resulted in a mean ground elevation bias of +56 cm and an absolute RMSE of 3.7 m. A more detailed description of how these parameters are used in the algorithm and how the noise-tracking is defined can be found in Hancock et al. (2011), Hancock et al. (2017) and Hancock et al. (2019). The waveforms were processed to remove noise, identify the ground and extract metrics that have been shown to be correlated with AGB (Duncanson et al., 2020; Hancock et al., 2019; Dubayah et al., 2020). As the GEDI ground finding algorithm had not been finalized, parameters were optimized by minimizing the difference between the GEDI and ALS derived ground elevations

(Hancock et al., 2019). GEDI's AGB models are data-driven from a global set of field and simulated GEDI waveforms, stratified by Plant Functional Type (PFT) and geographic domainIn their study (Duncanson et al., 2020). The US-wide GEDI AGB models from Duncanson et al. (2020) were applied to the simulated GEDI Level 2A Relative Height (RH) metrics (e.g. RH50, RH60, RH70, RH80, RH90, RH98 and their interaction terms) to obtain simulated AGB estimates at footprint level (GEDI-like Level4A) across the study area. In this study, these AGB models were fitted using Ordinary Least Squares (OLS) regression with exhaustive variable selection and showed R^2 ranging from 0.58 to 0.64 and relative RMSE ranging from 26.77% to 34.87% (for more details, see Duncanson et al., 2020 and Table S1). For the GEDI footprint products (e.g. Level 2A and Level 4A), it is expected that canopy penetration would be lowest using the coverage beam data acquired during the day. Moreover, these samples over dense canopies will likely be filtered because they may not reliably detect the ground in some situations. As such, we have filtered simulated coverage daytime beams from this analysis. Herein, we have not accounted for the GEDI footprint geolocation errors as our AGB estimates are aggregated at 1-ha for further modeling and mapping (Section 2.3).

2.2.3. Simulated AGB estimates from Ice, Cloud, and land Elevation Satellite 2 (ICESat-2)

NASA's ICESat-2 is a spaceborne lidar system primarily focused on solid earth and ice applications. Although not optimized for vegetation studies, preliminary research suggests it will be useful for forest height and potentially AGB estimation (Montesano et al., 2015; Swatantran et al., 2016; Duncanson et al., 2020; Neuenschwander and Pitts, 2019). The Advanced Topographic Laser Altimeter System (ATLAS) instrument on-board ICESat-2 is a photon counting laser altimeter. ATLAS operates at 532 nm at a rate of 10 kHz in the green range of the electromagnetic spectrum. The combination of the satellite velocity and laser repetition rate will result in one outgoing along-track laser pulse approximately every 70 cm on the Earth's surface, with each pulse illuminating a ~14 m diameter footprint. There are two geophysical data products specifically focused on land and vegetation: ATL08, the Level 3 along-track data product and ATL18, the Level 4 gridded product. ATL08 reports terrain height estimates, canopy height estimates, and relative height metrics such as: RH50 and RH98 over a 100 m transect.

As with the GEDI data, an ICESat-2 simulator was applied to the ALS data from Sonoma County, and height metrics were calculated for 100-m along-track transects simulating two photon counting rates (e.g. Neuenschwander and Pitts, 2019). We focus this analysis on simulations of ICESat-2 rather than on-orbit data for two reasons. First, at the time of writing very limited high quality (e.g. cloud-free) ICESat-2 data are available, thus the lidar sample would be very sparse and not representative of the capability of ICESat-2 for fusion with SAR. Using two-years of simulated data allows for a realistic spatial distribution of data for training SAR models. Secondly, there is a significant temporal difference between the time of field and airborne acquisitions used to generate the airborne lidar reference maps and the acquisition of on-orbit ICESat-2. Preliminary analysis of on-orbit ICESat-2 data over Sonoma County show a mean photon rate of 1.5 photons per shot with the low end approximately one photon per shot and the high end approximately two photons per shot. For this analysis, we simulated the ICESat-2 data over Sonoma at the low rate of one photon per shot, hereafter called the 'low photon rate', and one with two photons returned per shot, matching some on-orbit data in more optimal conditions with higher atmospheric transmission as found in many other ecosystems, hereafter called the 'high photon rate'. Herein, we used a noise rate of 0.1 MHz. These two sets of ICESat-2 simulations (with high and low return rates) were simulated to represent two years of coverage, and height metrics were calculated following the ICESat-2 ATL08 algorithm (Neuenschwander and Pitts, 2019). The same continental GEDI AGB models were applied to these ICESat-2 simulations and showed R^2 ranging from 0.60 to 0.69 and relative RMSE ranging from 27.02 to 29.22% (Duncanson et al., 2020). Although the spatial resolution of ICESat-2 height metrics are different from GEDI (100 m transect vs. ~25 m footprint) and are calculated differently (GEDI's relative heights are calculated including ground energy and ICESat-2 are calculated only for photons classified as canopy photons), the operational GEDI algorithm was useful for estimating AGB with ICESat-2 in this particular case (Duncanson et al., 2020). While further ICESat-2 AGB algorithm development will likely improve ICESat-2 based estimates, these early results show promise and serve as a useful input to develop GEDI ICESat-2 fused AGB products as explored in this paper.

2.2.4. NASA-Indian Space Research Organization (ISRO) Synthetic Aperture Radar (NISAR) simulations

NISAR is a spaceborne platform with both L-band (24 cm wavelength) and S-band (10 cm wavelength) radars. It will collect data within a spatial resolution of 7 m over a swath width ~ 242 km using a 12 m diameter deployable mesh reflector (NISAR User Handbook, 2016). The satellite will be in a 12-day repeat orbit at an altitude of 740 km (Rosen et al., 2016). NISAR will therefore achieve global coverage every 6 days considering both ascending and descending orbital passes. The currently planned NISAR acquisition strategy for co-polarization mode (HH) to be acquired every 6-days and cross-polarization mode (HV) 3 times every 24 days. This results in each area of the globe being imaged approximately ~60 times per year.

In this study, Uninhabited Airborne Vehicle Synthetic Aperture Radar (UAVSAR) data was used for simulating NISAR-like data over Sonoma. UAVSAR is an airborne fully polarimetric L-band (wavelength of 23.5 cm) synthetic aperture radar (Hensley et al., 2009) to provide robust repeat-pass radar interferometric measurements of deformation from both natural and anthropogenic sources. UAVSAR data were collected over Sonoma County on August 29th 2014 with a flight bearing of approximately 55°. Six flight-lines were acquired over Sonoma county. Each image is 20 km wide and approximately 155 km long, with nominal incidence angles for all swaths ranging from the near to far range between 21° to 65° and spatial resolution of 6.14 m (0.00005556 degrees at the equator), respectively. We extract the UAVSAR swath corresponding to NISAR to simulate NISAR. The data are distributed with 36 nominal looks which will be further averaged to generate the 1-ha AGB map. This high number of looks mimics the time-series averaging processing performed to generate the NISAR products. Thus no speckle filtering of the UAVSAR data is required. We note that UAVSAR is currently being used as part of the calibration and validation program for NISAR operational processing and science products.

Complex topography causes increased variation in SAR backscatter and limits its use over irregular terrain unless corrected for. Therefore, the UAVSAR imagery was radiometrically corrected for variations in illuminated area on the ground, as well as variations in forest reflectivity with viewing and terrain geometry, using the radar incidence angle and range slope, as in Simard et al. (2016). This algorithm takes into account aircraft attitude, antenna steering angle, and target geometry and can be applied to a variety of L-band backscatter images spanning a wide range of incidence angles over terrain with significant topography. It includes elements of both homomorphic and heteromorphic terrain corrections in order to correct for topographic effects and variations of canopy reflectivity with viewing and tree-terrain geometry (Simard et al., 2016). The NISAR processor assumes the steering angle to be zero. The radiometric and terrain correction was performed using the open-source software found at the GitHub repository (https://github.com/SimardJPL/UAVSAR-Radiometric-Calibration). Ideally, the RTC should be based on an L-band derived Digital Surface Model (DSM) to represent the L-band phase center. However, such DSM is not available. Instead, the ALS digital terrain model (DTM) was utilized to perform this radiometric correction. Assuming NISAR will use the best freely available DSM for radiometric calibration of backscatter images, we used the highest quality elevation model available: the ALS DTM. We compared the results obtained using the ALS DTM and the SRTM DSM, and found

the use of SRTM may cause a negative bias in aboveground biomass bias below 2 Mg/ha (<1%) in these mountainous regions. The 15-year time span since SRTM (year 2000) may also carry changes in canopy structure. These differences, in particular the small scale variations in local slopes, are further reduced by the averaging process to generate the 1-ha AGB products. The radiometric and geometrically corrected UAVSAR imagery was re-projected to UTM (WGS 1984 UTM Zone 10 N) and resampled to 30 m pixels to match the ALS-derived AGB map, using bilinear interpolation. Lastly, the NISAR-like data was simulated selecting only UAVSAR data within a nominal incidence angle ranging from 33° to 47° and slopes ranging from −10° to 10°.

2.3. Aboveground biomass modeling and assessment

Herein, we introduce our data fusion framework to calibrate a regional L-band AGB model using sparse estimates from GEDI and ICESat-2 to obtain wall-to-wall estimates. We develop two variations on this approach according to spatial segmentations of our L-band imagery in which our estimations occur. We consider (a) the segmentation along 1-ha regular square grid cells and (b) the irregularly shaped object-based segmentation outlined in (Clewley et al., 2014) and implemented using RSGISLib software (Bunting et al., 2014).The segmentation uses a K-means algorithm with iterative elimination to cluster locally homogeneous areas of backscatter of size equal to or greater than 1-ha. These approaches are subsequently referred to as "grid-based" and "object-based" approaches respectively.

First, each segment is populated with GEDI and/or ICESat-2 estimated AGB whenever lidar samples fall within a segment. When there are multiple samples from GEDI and/or ICESat-2 within a particular segment, we populate the segment according to a weighted average determined by the relative size of the lidar footprint. Specifically, ICESat-2's footprint, determined from a 100 m × 14 m rectangle, is ~3.5 larger than GEDI's, which is determined from a ~ 25 m circular footprint. Additionally, we populate all segments with the mean HV γ^0 backscatter in decibels (dB). Those segments with GEDI and/or ICESat-2 samples are used to calibrate our L-band AGB model. Second, we model the relationships between backscatter and AGB using the empirical model from (e.g. Yu and Saatchi, 2016; Mitchard et al., 2009; Schlund et al., 2018; Santoro and Cartus, 2018; Lucas et al., 2010):

$$HV\ Backscatter\ (dB) = \beta_0 + \beta_1 * log\left[AGB\left(\frac{Mg}{ha}\right)\right] + \varepsilon \text{ with } \varepsilon \tilde{} N(0, \sigma^2) \quad (1)$$

We fit Eq. (1) using binned AGB values with a 5 Mg/ha bin size and average the corresponding HV backscatter within each bin in order to reduce the HV backscatter noise and allow the underlying correlation between backscatter and AGB (Yu and Saatchi, 2016; Duncanson et al., 2020). We fit the standard linear estimator associated with Eq. (1) and our binned calibration data using (Core Team, 2018). To estimate AGB from the measured backscatter, we invert Eq. (1) to obtain:

$$AGB\ (Mg/ha) = exp\left(\frac{HV\ Backscatter - \beta_0}{\beta_1}\right) \quad (2)$$

We compare our AGB estimates within a segment to the mean AGB according to our reference ALS-derived AGB estimates in the same segment. The L-band AGB model estimates are then evaluated in terms of coefficient of determination (R^2) (Eq. (3)), absolute Root Mean Square Error (RMSE; Mg/ha) (Eq. (4)) and mean difference (MD; Mg/ha) (Eq. (5)):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \widehat{y}_i\right)^2}{\sum_{i=1}^{n}(y_i - \bar{y})} \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(\widehat{y}_i - y_i\right)^2}{n}} \quad (4)$$

$$MD = \frac{1}{n}\sum_{i=1}^{n}\left(\widehat{y}_i - y_i\right) \quad (5)$$

Above, y_i is the reference AGB in segment i, $\widehat{y}_i$ is our estimated AGB in segment i, and $\bar{y}$ is the mean biomass with respect to the ALS biomass reference map. We additionally calculate the relative RMSE and MD by dividing the respective absolute values (Eqs. 4 and 5) by the mean of GEDI and ICESat-2 AGB over our study area.

2.4. Regional aboveground biomass estimates

Regional estimates of AGB stocks were then computed by applying Eq. (2) at the landscape level (See Fig. 2). R^2 (Eq. (3)), RMSE (%) (Eq. (4)) and MD (%) (Eq. (5)) were computed from the linear relationship between estimated and reference AGB stocks at the segment scale for assessing the accuracy of AGB estimates at the regional scale. Also, we compared the regional estimates across fusion data scenarios using the so-called Wilcoxon–Mann–Whitney rank-sum (W) test (Wilcoxon, 1945) to assess if the average of the regional AGB estimates derived from this study differed from the average of the reference ALS map at a significance level of 5%. The difference between AGB estimates and the reference AGB map at the segment scale was then mapped.

We then compared the following five scenarios, all of which use NISAR data: models calibrated with only (1) GEDI or ICESat-2 data using high (2) and low (3) photon rates, and models calibrated with both (4) GEDI and ICESat-2 (high) and (5) GEDI and ICESat-2 (low) (Table 1).

3. Results

3.1. ALS and pre-calibration GEDI, ICESat-2 l and ICESat-2 h aboveground biomass estimates

The ALS and pre-calibration GEDI, ICESat-2l and ICESat-2h derived AGB estimates are presented in Fig. 3 and Fig. 4 at their original resolutions and then aggregated at grid and object segments. We observed significant differences among the average of the AGB from GEDI, ICESat-2l and ICESat-2h and when compared to the reference ALS-derived AGB average ($W \leq 0.35 \times 105$, p-value <0.05) (Fig. 3). In general, GEDI, ICEsat-2l and ICEsat-2h and their combinations, underestimated AGB when compared to ALS (Fig. 3 a-c).

The number of GEDI, ICESat2l and ICESat2h observations falling in the study area were 3876, 882 and 1943, and occupied 9.7, 4.4 and 8.9% of the total number of grids (n = 17,433) and 15.8, 7.9 and 15.6% of the total number of segments (n = 8777). After combining the GEDI and ICESat-2 derivations, the segments occupied by GEDI+ICEsat-2 l and GEDI+ICEsat-2h samples increased to 13.76% and 17.91% for the grid approach and 22.23% and 28.72% for the object approach, respectively (Fig. 4). In general, the number of GEDI, ICESat-2l and ICESat-2h observations falling within 1-ha-grid cells or 1-ha segment objects over the simulated two years of data acquisition were higher than 10 observations in some areas. But, as segments cover an equal or larger area then grid objects, the numbers of GEDI and ICESat-2 observations were higher within objects than grid cells (Fig. 5), thus reducing the uncertainty of the object-oriented approach with respect to the gridded approach. The majority of the grid cells had ~1–2 GEDI and ICESat-2 observations while the majority of objects had 1–4 GEDI and ICESat-2 observations. Combining both GEDI and ICESat-2 increased the number of samples per grid and segment, and significantly changed the distribution of the number of samples per grid or object compared to ICESat-2, but only slightly when compared to GEDI (Fig. 5c). The average footprint observations per grid and object were 1.31 (SD ±

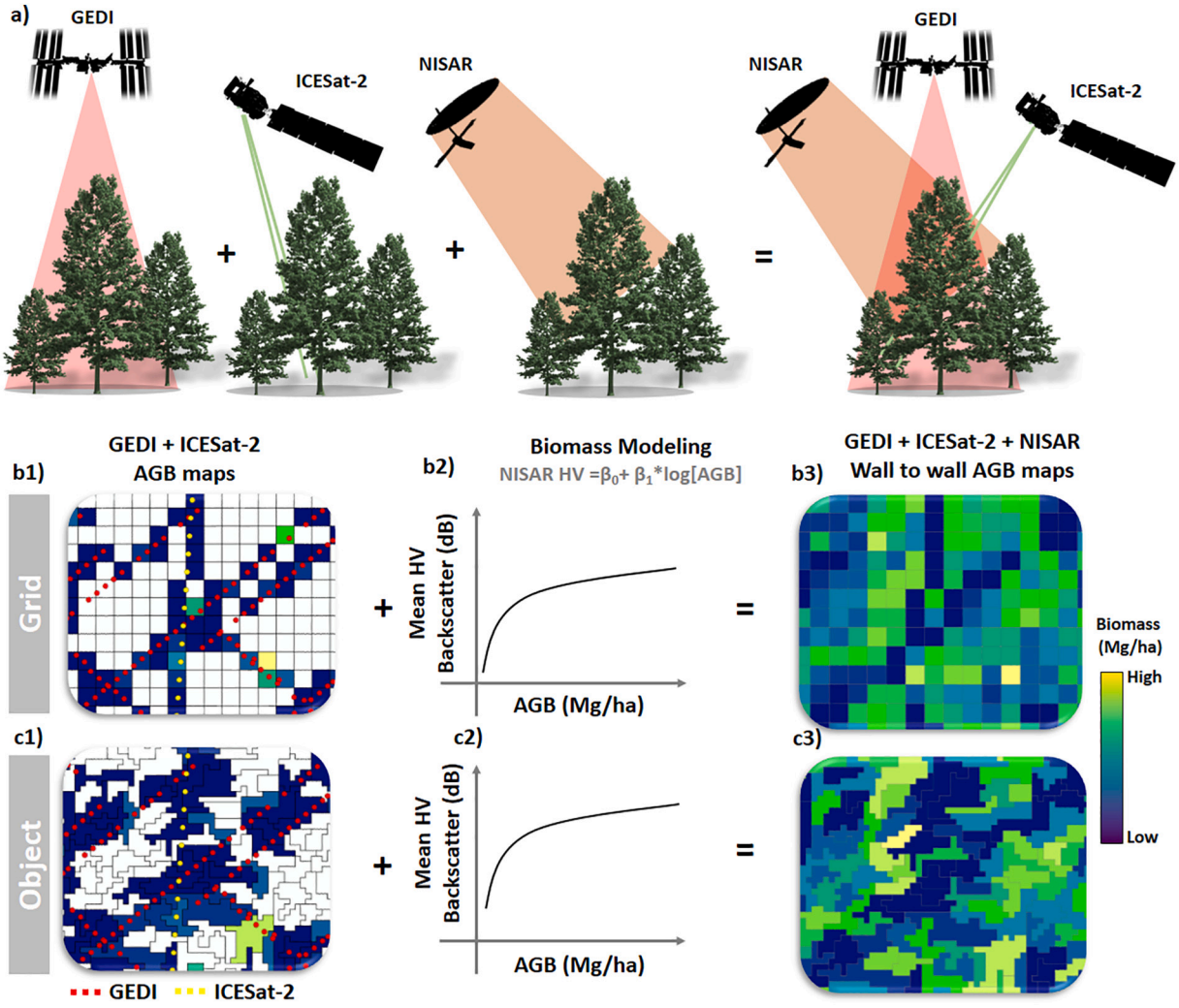

Fig. 2. Simplified illustration of the two GEDI, ICESat-2 and NISAR data fusion approaches for wall-to-wall AGB mapping. GEDI, ICESat-2 and NISAR missions a); GEDI and ICESat-2 fused AGB calibration at regular grid-based (b1) and object-based segments (c1); AGB modeling from NISAR-like HV backscatter (b2) and (c2); wall-to-wall AGB maps derived from the fusion of GEDI, ICESat-2 and NISAR datasets with grid-based (b3) and object-based segments (c3).

Table 1
GEDI, ICESat-2 and NISAR fusion scenarios for wall-to-wall AGB mapping.

Scenarios	Fusion description
GEDI +NISAR	GEDI data acquired at day and nighttime using power beam and at night using cover beam + NISAR HV Backscatter
ICESat-2 l + NISAR	ICESat-2 data acquired at day and nighttime using low photon return rate + NISAR HV Backscatter
ICESat-2 h + NISAR	ICESat-2 data acquired at day and nighttime using high photon return rate + NISAR HV Backscatter
GEDI+ICESat-2 l + NISAR	GEDI data acquired at day and nighttime using power beam and at night using cover beam + ICESat-2 data acquired at day and nighttime using low photon return rate + NISAR HV Backscatter
GEDI+ICESat-2 h + NISAR	GEDI data acquired at day and nighttime using power beam and at night using cover beam + ICESat-2 data acquired at day and nighttime using high photon return rate + NISAR HV Backscatter

0.56) and 1.65 (SD ± 1.16) when combining GEDI and ICESat-2l estimates, and 1.33 (SD ± 0.58) and 1.71 (SD ± 1.19) when combining GEDI and ICESat-2h estimates, respectively.

3.2. Aboveground biomass models

For the five scenarios described in Table 1, we report in Table 2 the results of the model accuracy for estimating AGB with both grid and object-based segmentation approaches. For all models, HV backscatter explained more than 60% of variations in AGB with relative RMSE ≤ 78.27% (130.57 Mg/ha) and MD ≤ + − 39.09% (−67.58 Mg/ha), respectively (Table 2). The positive values in MD indicate that the models are slightly overestimating AGB while negative values indicate underestimating on our calibration segments. The L-band models calibrated with both simulated GEDI and ICESat-2 data had slightly higher accuracy than those calibrated with only GEDI or ICESat-2 alone. Also, L-band AGB models calibrated with ICESat-2l provided lower accuracy than those models calibrated with ICESat-2h. In general, our fusion framework applied to a grid-based segmentation of the scene showed no significant difference in terms of R^2, relative RMSE and MD than when the same framework was applied to the object-based segmentation (Table 2; Table S2).

3.3. Wall-to-wall regional aboveground biomass estimates derived from the fusion framework

Wall-to-wall AGB estimates, such as those derived from GEDI + ICESat-2h + NISAR scenarios, both using grid and object-based fusion

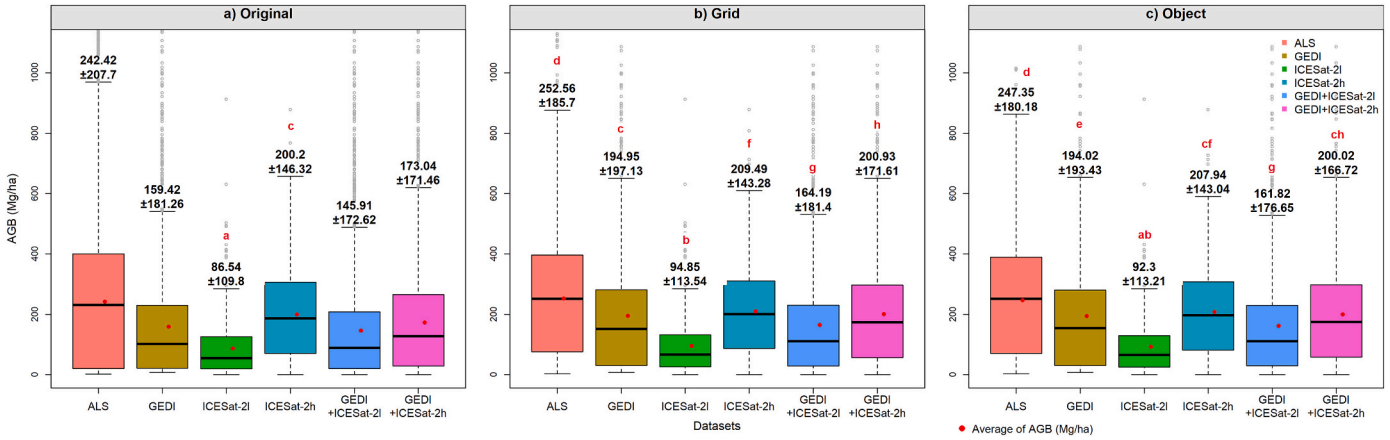

Fig. 3. ALS and GEDI, ICESat-2l and ICESat-2h derived AGB (Mg/ha) estimates at original resolutions (a) and then aggregated at grid (b) and object (c) levels for the entire study area. Wilcoxon–Mann–Whitney rank sum test results are shown by the letters in red. Datasets that share the same letter do not differ statistically in terms of average of AGB at a significance level of 5% according to the Wilcoxon–Mann–Whitney rank sum (W) test. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

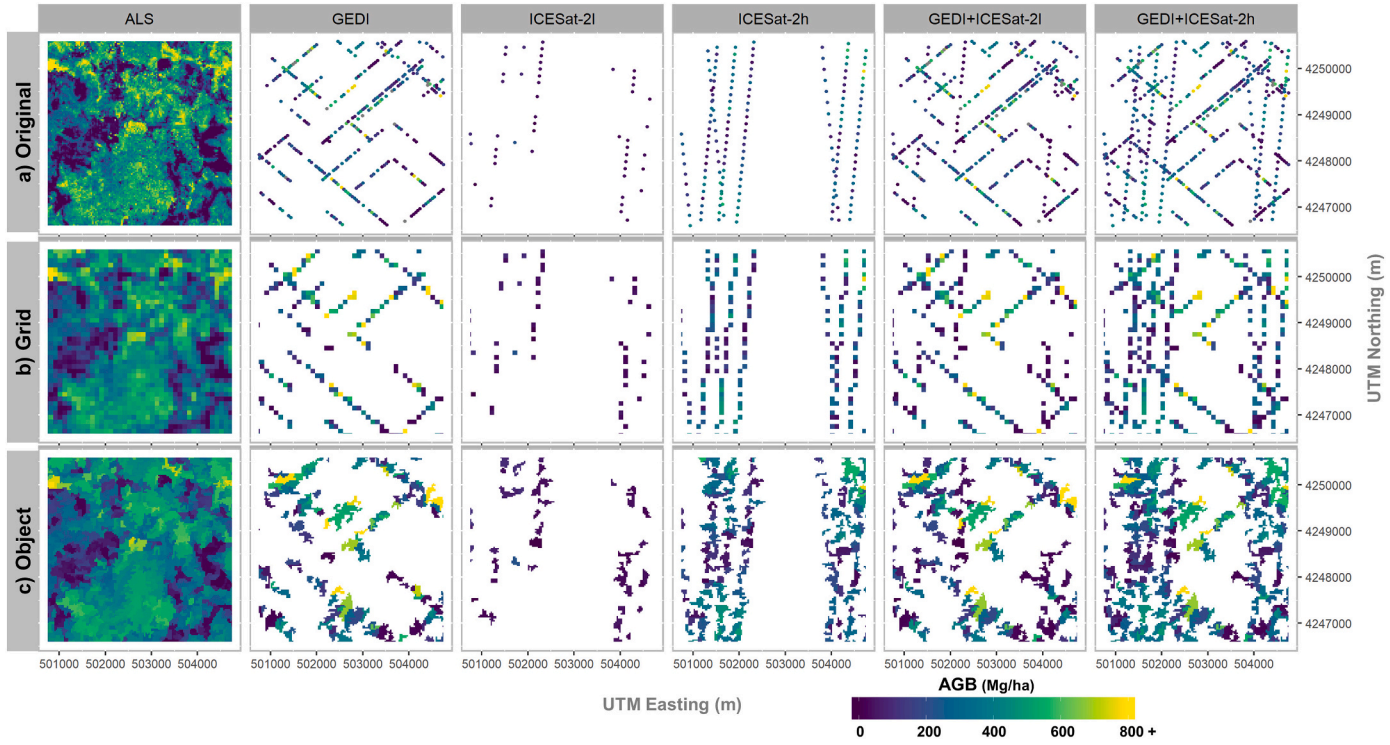

Fig. 4. Illustration of ALS and GEDI, ICESat-2l and ICESat-2h derived AGB (Mg/ha) estimates over two years of data acquisition at original resolutions (a) and then aggregated at grid (b) and object levels (c) for a subset of study area (1600 ha).

data approaches, showed closer proximity to the reference ALS AGB map compared to those scenarios use only GEDI or ICESat-2 combined with NIAR. ICESat-2l strongly underestimated AGB estimates when compared to reference ALS-derived AGB estimates using both grid and object-based fusion data approaches (Fig. 6). When comparing the wall-to-wall AGB maps derived in this study (Fig. 7) against the ALS-derived AGB map at the segment level, the combination of GEDI+ICESat2 + NISAR produced AGB maps with the highest accuracies and lowest average differences (Fig. 7 and Fig. 8) than those scenarios that use only the combination of GEDI or ICESat-2 with NISAR data (Fig. 7 and Fig. 8). We highlight that for all scenarios and segmentations except models calibrated with ICESat-2l, we obtain RMSE below ≤57.17% and MD ≤ 15.16% (Fig. 8) when comparing the AGB estimates to the reference ALS-derived AGB map.

4. Discussion

This paper focused on developing and testing fusion strategies to combine the sparse spaceborne lidar AGB estimates from GEDI and ICESat-2 with wall-to-wall L-band SAR data with an eye toward fusion of on-orbit data from the spaceborne lidars and spaceborne SAR data such as from NASA/ISRO's NISAR. While the approaches tested here do not represent the breadth of possible methods for combining these mission datasets, they align with the general concept of training wall-to-wall data with lidar samples as in Lefsky (2010), Qi et al. (2019a) and Saarela et al. (2018).

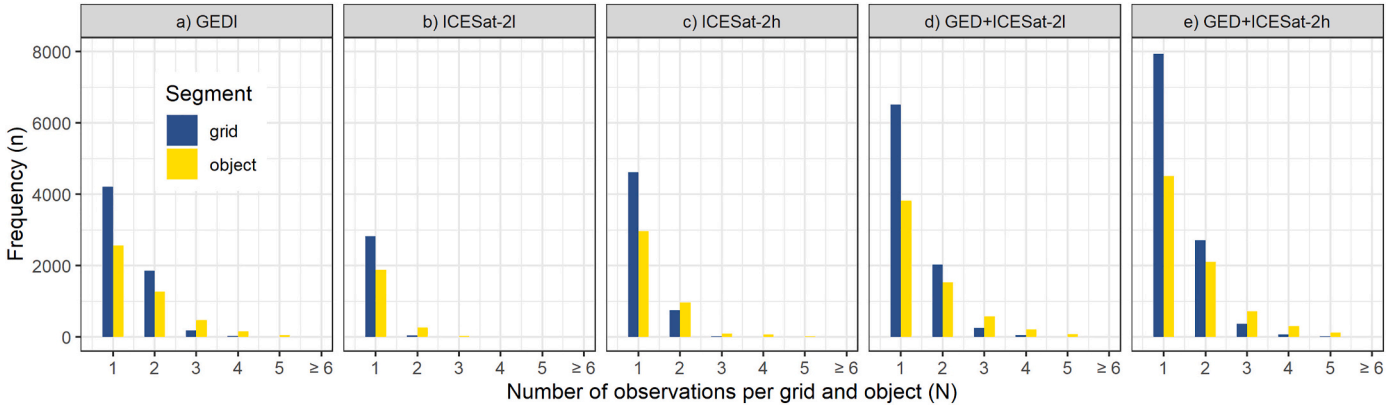

Fig. 5. Density plot of the number of GEDI (a), ICESat-2 (b) and combined GEDI and ICESat-2 (c) observations (footprints) per grid (red) and objects (green) approaches across the study area simulated over two years of data acquisition. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Accuracy and model parameters of the fitted NISAR-like AGB models across scenarios and segmentation approaches. R^2, RMSE and MD are evaluated on the segments used to calibrate the L-band model. SE is the standard error of R^2.

Fusion Approaches	Scenarios	Model		$R^2 \pm$ SE	RMSE		MD	
		β_0	β_1		Mg/ha	%	Mg/ha	%
Grid	GEDI+NISAR	−32.19	3.57	0.64 ± 0.02	113.84	53.61	−7.80	−3.68
	ICESat-2l + NISAR	−34.6	4.45	0.63 ± 0.03	130.57	78.91	−62.94	−38.04
	ICESat-2h + NISAR	−36.09	4.26	0.61 ± 0.02	115.86	49.93	−16.25	−7.01
	GEDI+ ICESat-2l + NISAR	−30.79	3.33	0.63 ± 0.02	110.95	55.92	−5.95	−3.00
	GEDI+ ICESat-2h + NISAR	−32.75	3.62	0.62 ± 0.01	115.52	52.17	3.78	1.71
Object	GEDI+NISAR	−33.3	3.72	0.65 ± 0.02	113.84	51.41	−13.00	−5.87
	ICESat-2l + NISAR	−35.13	4.54	0.62 ± 0.03	133.87	77.44	−67.58	−39.09
	ICESat-2h + NISAR	−35.79	4.19	0.60.0.2±	119.43	49.86	−15.51	−6.48
	GEDI+ ICESat-2l + NISAR	−31.55	3.48	0.62 ± 0.02	114.10	54.82	−13.75	−6.61
	GEDI+ ICESat-2h + NISAR	−33.30	3.72	0.61 ± 0.02	119.20	51.87	−1.60	−0.706

Here, we use realistic proxies of on-orbit mission estimates to illuminate the importance of selection of lidar training samples, the utility of binning lidar AGB estimates when fitting backscatter models, and the similarity of gridded vs. object-based approaches. Note that we compare the estimates from each fusion strategy at a segment-level to the reference ALS map, but do not compare estimates of the uncertainties from each fusion strategy. Indeed, we do not propagate error from field estimates, simulated height metric uncertainty or model fits through to our mapped estimates, and therefore limit our analysis to comparison of model-based estimates from various combinations of lidar training samples in either grids or objects. The consistency found between these approaches suggests that the general method for training binned backscatter models with reliable lidar-derived heights is robust and not highly sensitive to the lidar mission used (provided height estimates are accurate), or the geometry of application (gridded vs. object-based). Indeed, the transferability of GEDI's algorithms to ICESat-2 is promising for the fusion of these lidar data streams toward larger training samples and better geographic coverage of mapped AGB products.

4.1. Considerations of selection of lidar calibration samples

The accuracy of any lidar-SAR fusion product that relies on calibration with lidar AGB samples will naturally depend on the accuracy of those samples. This study compared various combinations of simulated GEDI and ICESat-2 data produced by Duncanson et al. (2020) and found that the selection of lidar training data was a primary driver of fusion product utility. Most notably, accuracies were sensitive to the photon return rate for ICESat-2, with AGB estimates low photon return ICESat-2 data yielding large deviations from our AGB reference map. Selecting the highest quality lidar training data (i.e. high photon return ICESat-2 data) produced the highest accuracies.

We note that our fusion framework has not performed as well when calibrating with only ICESat-2l (Table 2 and Fig. 6), i.e. ICESat-2 data with a "low photon return" rate (Section 2.2.3). We conjecture that for accurate AGB calibration using ICESat-2 AGB estimates from operational GEDI models in complex forest ecosystems, the samples for calibration must have a high photon return rate. At this time it is uncertain whether ICESat-2's lower photon rates would be the result of atmospheric conditions affecting the return signal or the complex forest and topographic structures of the target. Regardless, we recommend caution when using low photon return rate ICESat-2 data and recommend further analysis in different ecosystems to determine optimal thresholds for using ICESat-2 for forest structure mapping. Where higher photon rates are available, as have been observed in other areas of the world with potentially less atmospheric contamination (e.g. in Finland, Neuenschwander and Magruder, 2019) we expect ICESat-2 will provide an important secondary lidar dataset that would yield higher mapping accuracies when used in combination with GEDI than either mission would yield by themselves. We also note that in this research we were applying AGB models developed for GEDI to ICESat-2 height metrics (Duncanson et al., 2020). Further ICESat-2 biomass algorithm development and refinement will likely increase the accuracies of ICESat-2 estimates and thus fused products. Regardless, even here with GEDI adopted ICESat-2 AGB estimates the highest accuracies were found when using both simulated GEDI and ICESat-2 AGB samples to train backscatter models.

A second consideration is the number of training samples required per segment. We note that Silva et al. (2018) recommended a minimum of three large footprint lidar observations to provide an unbiased estimate of AGB within a 1-ha area in tropical forests. In this paper, where

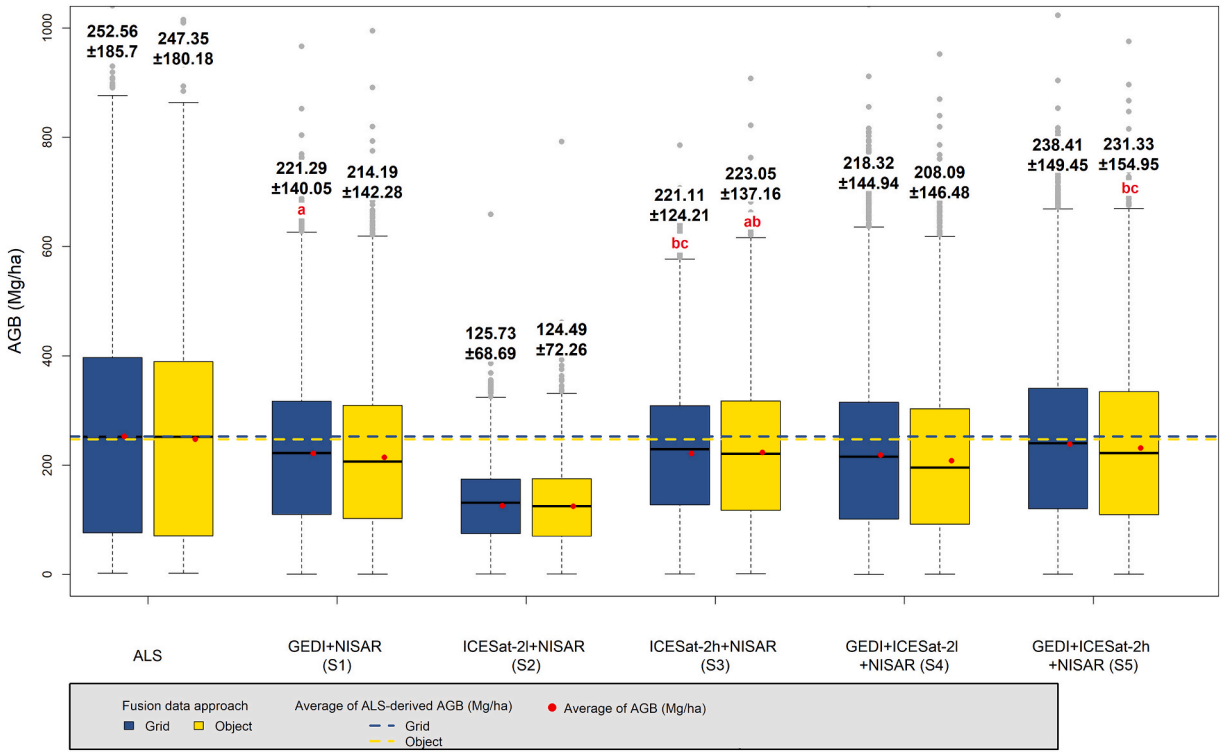

Fig. 6. Regional AGB estimates based on samples (a) and wall-to-wall maps (b) derived from grid and object-based segmentation approaches. Values shown in black at each boxplot represent the average ± standard deviation of AGB (Mg/ha) at the segment level for the entire study area. Wilcoxon–Mann–Whitney rank sum (W) test results are shown by the letters in red. Scenarios that share the same letter do not differ statistically in terms of average at a significance level of 5%. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

most segments had ~2 lidar samples, we were still able to estimate AGB accurately at the landscape scale. While it is uncertain whether this is because of the binning approach to fitting our L-band backscatter AGB models, or whether AGB in temperate forests is less spatially heterogeneous than tropical forests, we feel these results emphasize that relatively sparse lidar samples are useful for 1 ha AGB wall-to-wall mapping, and that the optimal shot density for training samples may vary by ecosystem.

We found an underestimation of AGB at footprint level in the highest AGB forests from the two simulated spaceborne datasets when compared to the ALS reference map. The ALS reference map also has uncertainties, but because it uses airborne lidar locally trained with field plots, it is of higher accuracy than the simulated satellite estimates. A thorough discussion of the observed underestimation is provided in Duncanson et al. (2020), but briefly this is attributed to well documented challenges of AGB estimation in high AGB forests from a combination of height measurement error of the sensors, but more importantly uncertainties in the models for estimating AGB in the field and from space. The allometric models applied to estimate AGB in the field have uncertainties that increase with tree size, and likely have biases in these ecosystems (Disney et al., in press). Similarly, the models applied to estimate AGB from GEDI and ICESat-2 are developed at a continental scale where most training data are from lower AGB forests. Therefore, this underestimation is likely attributed to the models applied to the GEDI and ICESat-2 lidar estimates, which include uncertainties in the field estimates. The purpose of this paper is to develop methods for fusing lidar estimates and SAR data for AGB mapping, and does not overcome challenges in lidar-based estimation in high AGB forests.

4.2. Backscatter models

Radar backscatter has been used for deriving AGB estimates in many forest ecosystems (e.g. Mitchard et al., 2009; Saatchi et al., 2011; Yu and Saatchi (2016); Schlund et al., 2018). Previous authors have discussed important factors that can affect the relationship between SAR backscatter and forest AGB. For instance, Yu and Saatchi (2016) pointed out that noise levels are high when correlating AGB directly with SAR backscatter, and many sources of error can contribute to this high noise at the pixel level. In this study, the NISAR-AGB models were fit using a binning strategy (Yu and Saatchi, 2016), in addition, to a high number of looks and spatial averaging, which combined significantly reduce the effect of the backscatter noise on the relationship between HV backscatter and AGB. Also, due to the use of the analytical biomass model, our models naturally saturated at high AGB ranges, however, they were still sensitive to higher biomass regimes, as compared to those presented in Yu and Saatchi (2016).

A direct comparison of our AGB estimates over Sonoma to other SAR-AGB studies is difficult as each study is highly dependent on the sensors used and site of consideration. Moreover, a majority of previous SAR-AGB studies rely on a relatively small number of field measurements to calibrate and validate AGB estimation framework (Mitchard et al., 2009; Persson and Fransson, 2014; Shao and Zhang, 2016; Schlund et al., 2018; Yu and Saatchi, 2016; Sinha et al., 2015). In this work, we provide a different kind of validation for our methodology as we use an ALS-derived AGB map. This affords orders of magnitude more validation and calibration data than those studies validating their estimates on field measurements. Also, this validation allows us to evaluate how well we were able to extend the lidar AGB estimates across our region of interest with L-band SAR. In contrast to numerous studies, we consider forest areas with significantly higher AGB density and consequently more complex canopy structure than those studies that focus on areas with sparser canopy and lower biomass density (Sinha et al., 2015). For instance, in Neumann et al. (2012), the authors successfully used L-band SAR for estimating AGB over the boreal forest in Québec. However, AGB range of their boreal region of interest is bound by 200 Mg/ha. Furthermore, their approach requires fully polarimetric radar which

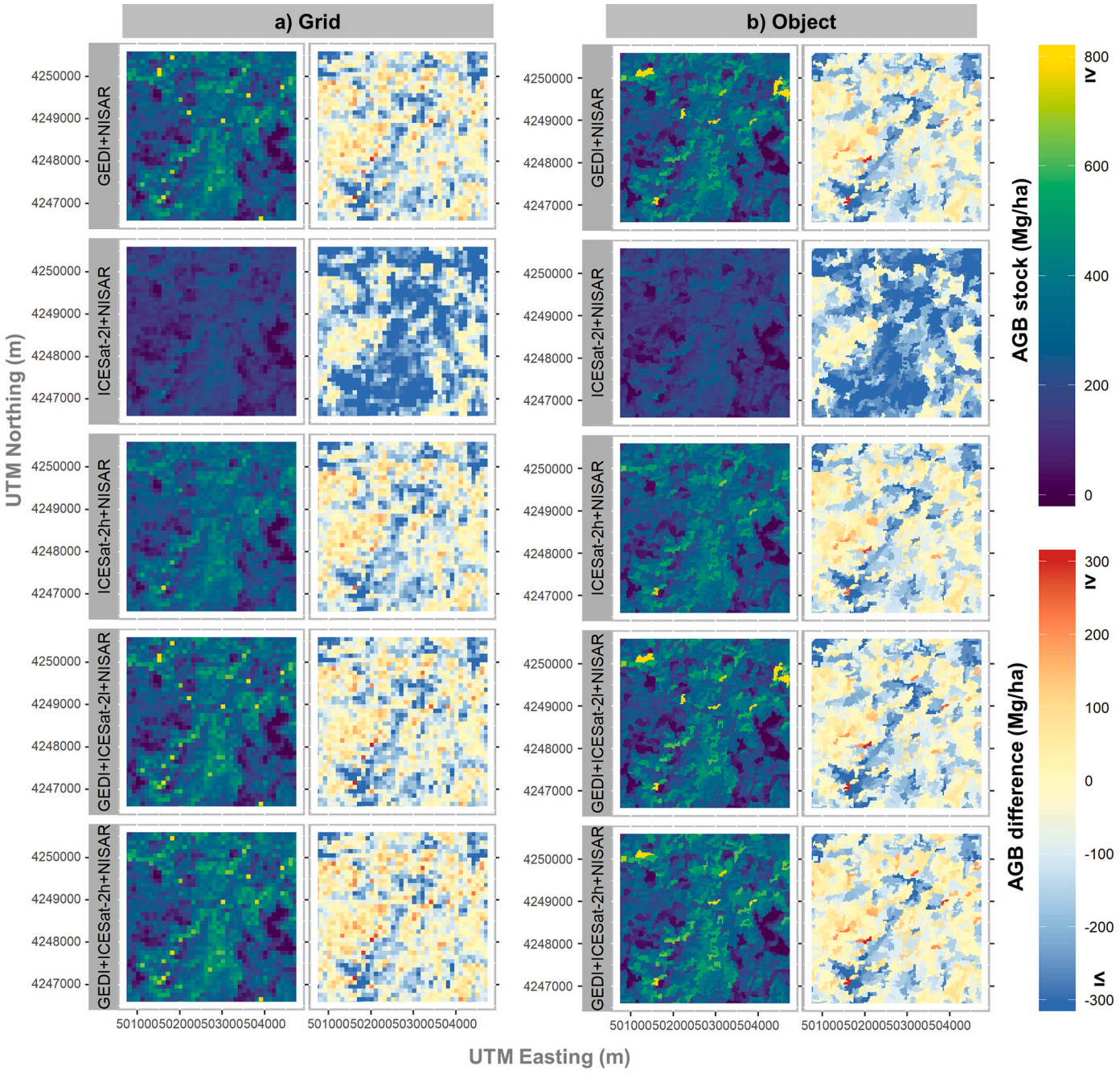

Fig. 7. Illustration of ALS and NISAR L-band derived wall-to-wall aboveground biomass (AGB) maps as results of our fusion framework. (a) grid and (b) object-based approaches showed in a subset scale (1600 ha). Each column represents the AGB estimates and differences from the proposed fusion framework. Each row represents one of the five calibration scenarios from Table 1. AGB difference maps are computed with respect to the ALS reference map.

may not be globally available with NISAR. We highlight that our approach requires only HV backscatter and will be globally applicable with NISAR. In Bouvet et al. (2018), they obtain RMSE below 20 Mg/ha, but again only consider areas of low AGB density, namely areas with pixelwise AGB bounded by 100 Mg/ha. They use dual polarization HH + HV backscatter and so will be valid with NISAR data (NISAR handbook). Indeed, the authors applied their model to the entire continent of Africa using ALOS-1 mosaics from (Shimada et al., 2010).

For practical use, in contrast to Silveira et al. (2019), we have not found significant differences between the AGB estimation accuracies when using grid or object-based approaches for AGB mapping from GEDI, ICESat-2 and NISAR. However, a direct comparison of the AGB estimates from these two segmentations is difficult as each spatial segmentation requires a different validation set. Specifically, the validation set for a particular segmentation is determined by populating each segment with the mean biomass from our ALS-derived reference biomass map. The object-based segments are variably sized clustering areas of similar backscatter likely preserving the range of AGB values originally present in the AGB reference map. In contrast, the grid-cell aggregates spatially nearby pixels without reference to backscatter thus smoothing the distribution of AGB in the populated reference AGB product. We highlight that R^2, RMSE and MD are similar for both segmentations even with these differences in the validation set. This appears to indicate that having a minimum size for irregular-based object-based segments that match the desired spatial resolution of grid-based segments will produce similarly accurate biomass maps. We add that

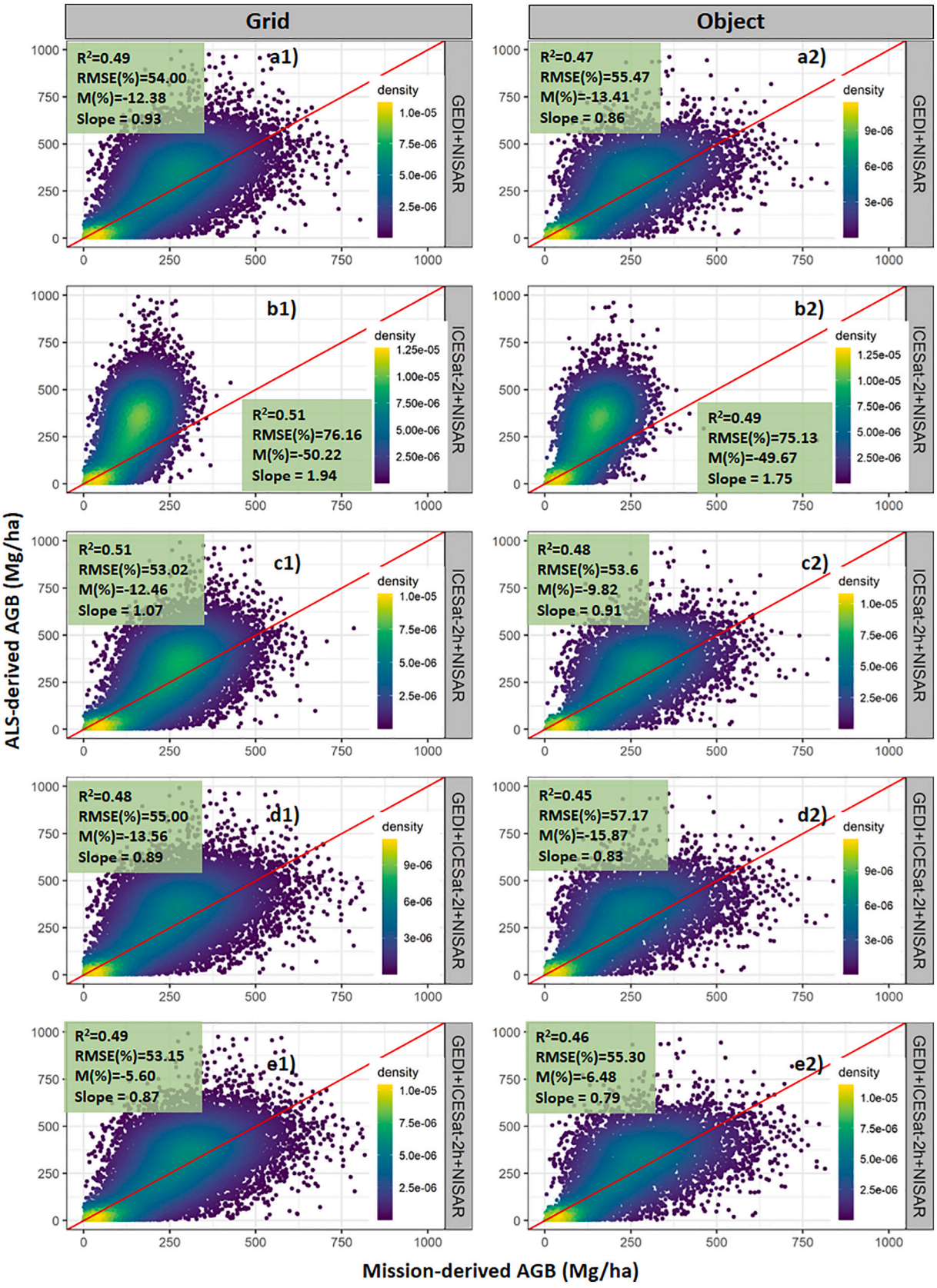

Fig. 8. ALS and NISAR L-band derived wall-to-wall regional aboveground biomass (AGB) estimates using grid (first column) and objects (second column) based approaches across calibration scenarios (rows). The red solid line represents the 1:1 relationship. R^2, RMSE and MD are computed at the segment level (Eqs. (3)–(5)) by the linear relationship between estimated (output from our fusion framework) and reference ALS-derived AGB. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

object-based spatial segmentations have an associated computational cost, both to retrieve the initial segmentation as well as to populate the relevant values for calibration and subsequent biomass mapping. Ultimately, each spatial segmentation provides slightly different advantages and trade-offs that must be weighed for the given application and study area.

We expect more lidar samples per calibration segment to mitigate noise in the lidar biomass estimates and increase the fidelity of the final biomass estimate within a calibration segment, but limited by accuracy of the lidar estimates themselves, and their freedom from systematic biases. Increasing the spatial segmentation sizes must be done with caution as a larger segment can have greater biomass heterogeneity and possibly introduce sampling error in our weighted averaging if the segment's size far exceeds the size of the lidar footprint. Because the object-based segments are better able to capture homogeneous biomass areas than grid-based segments, coarsening the spatial resolution with object-based segments may mitigate possible sampling error. Generally, the size of the spatial segmentation must be further assessed to determine how to optimally map biomass. Such considerations will vary from region to region depending both on the heterogeneity of biomass in that particular area as well as the total lidar samples available. One important line of inquiry here may be to base such segmentations on structural typologies derived from fusion of lidar and SAR, separate from AGB biomass estimation to provide physically-based foundations for the segmentation.

4.3. Considerations for wide area biomass mapping

Demands for accurate, high-resolution, wall-to-wall maps of forest AGB are increasing (Glen et al., 2016). Fusion of multiple remote sensing data sources, such as GEDI, ICESat-2 and NISAR present an exciting opportunity for achieving a desired high level of accuracy at user-relevant resolutions for wide-area AGB mapping. In this study, our fusion framework leverages the lidar training samples and their underlying AGB models, while taking advantage of the wall-to-wall SAR data to produce more accurate regional mean estimates than using the lidar data alone. The deviation between the reference map mean AGB density and the simulated spaceborne estimates was only 16.02 Mg/ha (6.48%) when using both the grid and object-based approaches. This improvement in regional estimation over the lidar results suggests that sampling error may contribute to errors in lidar-based regional biomass estimation in this high biomass, heterogeneous landscape, but that the use of wall-to-wall data for extrapolation of lidar estimates is a powerful path forward to fill the increasing demand for AGB products.

While we tested the NISAR L-band backscatter for mapping GEDI and ICESat-2 AGB-derived estimates across a landscape, other remote sensing data, such as Tandem-X (e.g. Qi et al., 2019a) and Tandem-L (e.g. Krieger et al., 2010; Moreira et al., 2018) InSAR, ALOS-2 PALSAR-2 (e.g. Marshak et al., 2019) and Landsat 8 Operational Land Imager (OLI) (e.g. Li et al., 2020) may also be used. For instance, Qi et al. (2019a) suggest the use of ancillary topographic data, such as those produced from GEDI, to reduce biases to acceptable levels when estimating spatially continuous canopy structure with RVoG model and Tandem-X InSAR. Also, Glenn et al. (2016) using a combination of ICESat-2 and Landsat 8 OLI demonstrated potential synergies between these two sensors for spatially contiguous vegetation cover mapping. Moreover, even though in this study we focused on using simulated NISAR L-band backscatter for wall-to-wall AGB mapping using sparse GEDI and ICEsat-2 simulated data, we believe our results are also relevant to other missions dedicated to estimate AGB, such as the ESA BIOMASS (Carreiras et al., 2017; Quegan et al., 2019) and TanDEM-L (Krieger et al., 2010; Moreira et al., 2018).

Although, in the remote sensing research community, data fusion has played a significant role for decades (Schmitt and Zhu, 2016), careful attention to the unique characteristics of each input dataset is also critical for the effective application of fusion algorithms. In this study, many factors, including fusion methodology, data acquisition, NISAR HV backscatter segmentation, spatial resolution among GEDI, ICESat-2 and NISAR data, etc. impacted our reported accuracies of fused AGB maps. We derived wall-to-wall AGB maps using parametric models to relate lidar AGB estimates to backscatter, but the results could vary if we had used other modeling approaches or other ancillary data. For instance, improvements on AGB estimation have been found when combined lidar, satellite radar and multispectral imagery data with machine learning algorithms, such as Random Forest (Breiman, 2001) and *k*-NN Imputation (Crookston and Finley, 2008) for vegetation mapping (Garcia et al., 2017; Yu and Saatchi, 2016). Therefore, further work involving machine learning approaches or refined algorithms for NISAR backscatter segmentation (e.g. superpixel Marshak et al. (2019)) might help further increase the spatial resolution and accuracy of wall to wall AGB maps when combined with GEDI, ICESat-2 and NISAR data.

5. Conclusion

In this paper, we investigated a regional data fusion framework to extend reliable GEDI and ICESat-2 AGB estimates with NISAR-like data to obtain an approximately 1-ha resolution wall-to-wall AGB map. We calibrate regional L-band AGB models with different calibration combinations and two possible spatial segmentations of the L-band imagery. We compared this approach with simulated GEDI, ICESat-2 and NISAR data over Sonoma County along with a high-quality lidar-derived AGB reference map. Using this reference data, we showed that most variations of our GEDI, ICESat-2 and NISAR fusion framework improved AGB estimates either with grid or objects segments, both at the 1-ha segment level and the landscape level, compared to those obtained by any single sensor alone. We note that the ICESat-2 AGB estimates were unreliable when the photon return rate was low and recommend further research into the utility of ICESat-2 for forest AGB mapping in different ecosystems. We expect in many areas that ICESat-2 will have enough samples with high photon return rate over a 2 year period to make our proposed framework suitable not only with GEDI, but also in boreal forests where GEDI data are not available. The regional framework and results from this study underscore the power of multi-sensor fusion. Continued attention toward improving fusion estimation frameworks for AGB and its errors should be a priority.

Credit author statement

CAS, LD, AN, SH, NT, LF, MC and CM: conceptualization, methodology, data processing and writing up. RD, JA, SL and MH: contributed to the interpretation, quality control and revisions of the manuscript. All authors read and approved the final manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This study was primarily funded through NASA's Carbon Monitoring System (CMS, grant 15-CMS15-0055). Amy Neuenschwander was supported through ICESat-2 (NASA HQ NNX17AG55G). This work was partly conducted by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration (NASA). Simard and Marshak were supported by the Terrestrial Ecology Program. We thank Jim Kellner, Paul Patterson and Sean Healey for helpful input in the development of the ALS AGB models applied here. We also gratefully acknowledge Michael Falkowski, Andrew Hudak, Patrick Fakety, Warren Cohen, Hans Anderson, and Luigi Boscetti for provision of the US field and ALS data used to generate

the ALS algorithms in this study, and Suzanne Marselis and David Minor for data processing and database management. Additional thanks to Kaitlin Harbeck for helping produce the simulated ICESat-2 data, and Katherine Pitts for running the ATL08 algorithm on the simulated ICESat-2 data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2020.112234.

References

Berman, N., 2006. Sonoma County General Plan 2020. General Plan Update Draft Environmental Impact Report. Accessed on August 15, 2019. Available at. http://www.sonoma-county.org/prmd/docs/eir/gp2020deir/full_deir.pdf.

Bouvet, A., Mermoz, S., Le, T., Villard, L., Mathieu, R., Naidoo, L., Asner, G.P., 2018. Remote sensing of environment an above-ground biomass map of African savannahs and woodlands at 25 m resolution derived from ALOS PALSAR. Remote Sens. Environ. 206 (December 2017), 156–173. https://doi.org/10.1016/j.rse.2017.12.030.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Bunting, P., Clewley, D., Lucas, R.M., Gillingham, S., 2014. Computers & Geosciences the remote sensing and GIS software library (RSGISLib). Comput. Geosci. 62, 216–226. https://doi.org/10.1016/j.cageo.2013.08.007.

Carreiras, J.M.B., Quegan, S., Le Toan, T., Ho Tong Minh, D., Saatchi, S., Carvalhais, N., et al., 2017. Coverage of high biomass forests by the ESA BIOMASS mission under defense restrictions. Remote Sens. Environ. 196, 154–162. https://doi.org/10.1016/j.rse.2017.05.003.

Chen, Qi, McRoberts, Ronald E., Wang, Changwei, Radtke, Philip J., 2016. Forest aboveground biomass mapping and estimation across multiple spatial scales using model-based inference. Remote Sens. Environ. 184, 350–360. https://doi.org/10.1016/j.rse.2016.07.023.

Clewley, D., Bunting, P., Shepherd, J., Gillingham, S., Flood, N., Dymond, J., Lucas, R., et al., 2014. A python-based open source system for geographic object-based image analysis (Geobia) utilizing raster attribute tables. Remote Sens. 6 (7), 6111–6135. https://doi.org/10.3390/rs6076111.

R Core Team, 2018. R: A Language and Environment for Statistical Computing Vienna, Austria.

Crookston, N.L., Finley, A.O., 2008. yaImpute: an R package for kNN imputation. J. Stat. Softw. 23, 1–16. https://doi.org/10.18637/jss.v023.i10.

Deo, R.K., Russell, M.B., Domke, G.M., Andersen, H.-E., Cohen, W.B., Woodall, C.W., 2017. Evaluating site-specific and generic spatial models of aboveground forest biomass based on landsat time-series and LiDAR strip samples in the Eastern USA. Remote Sens. 9, 598. https://doi.org/10.3390/rs9060598.

Dubayah, R., Swatantran, A., Huang, W., Duncanson, L., Tang, H., Johnson, K., Hurtt, G. C., 2013. LiDAR-derived Biomass, Canopy Height and Cover, Sonoma County, California, 2013. CMS.

Dubayah, R., Bryan, J., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Silva, C., 2020. Science of remote sensing the global ecosystem dynamics investigation: high-resolution laser ranging of the earth's forests and topography. Sci. Rem. Sens. 1 (September 2019), 100002. https://doi.org/10.1016/j.srs.2020.100002.

Duncanson, L., Neuenschwander, A., Hancock, S., Thomas, N., Fatoyinbo, T., Simard, M., Flight, S., 2020. Remote Sensing of Environment Biomass estimation from simulated GEDI, ICESat-2 and NISAR across environmental gradients in Sonoma County, California. Remote Sens. Environ. 242 (April), 111779. https://doi.org/10.1016/j.rse.2020.111779.

FAO, 2018. The State of the World's Forests 2018 - Forest Pathways to Sustainable Development. Rome. License: CC BY-NC-SA 3.0 IGO.. Accessed on August 15, 2019. Available at. http://www.fao.org/3/I9535EN/i9535en.pdf>.

Garcia, Mariano, Saatchi, Sassan, Casas, Angeles, Koltunov, Alexander, Ustin, Susan, Ramirez, Carlos, Jorge Garcia-Gutierrez, H.B., 2017. J. Geophys. Res. Biogeosci. https://doi.org/10.1002/2015JG003315.

Glenn, N.F., Neuenschwander, A., Vierling, L.A., Spaete, L., Li, A., Shinneman, D.J., Mcilroy, S.K., 2016. Remote Sensing of Environment Landsat 8 and ICESat-2: performance and potential synergies for quantifying dryland ecosystem vegetation cover and biomass. Remote Sensing of Environment 185, 233–242. https://doi.org/10.1016/j.rse.2016.02.039.

Godfrey, S., 2015. Sonoma County Soil Types. Accessed on August 15, 2019. Available at <https://www.arcgis.com/home/item.html?id=9931ee57cbed4a0f9ed38505ae9f538a>.

Hancock, S., Disney, M., Muller, J.P., Lewis, P., Foster, M., 2011. A threshold insensitive method for locating the forest canopy top with waveform lidar. Remote Sens. Environ. 115 (12), 3286–3297. https://doi.org/10.1016/j.rse.2011.07.012.

Hancock, S., Anderson, K., Disney, M., Gaston, K.J., 2017. Measurement of fine-spatial-resolution 3D vegetation structure with airborne waveform lidar: calibration and validation with voxelised terrestrial lidar. Remote Sens. Environ. 188, 37–50. https://doi.org/10.1016/j.rse.2016.10.041.

Hancock, S., Armston, J., Hofton, M., Sun, X., 2019. The GEDI Simulator: A Large-Footprint Waveform Lidar Simulator for Calibration and Validation of Spaceborne Missions, pp. 294–310. https://doi.org/10.1029/2018EA000506.

Hensley, S., Zebker, H., Jones, C., Michel, T., Muellerschoen, R., Chapman, B., 2009. First deformation results using the NASA/JPL UAVSAR instrument. In: 2nd Asian-Pacific Conference on Synthetic Aperture Radar. IEEE, Xi'an Shanxi, China, pp. 1051–1055.

Hese, S., Lucht, W., Schmullius, C., Barnsley, M., Dubayah, R., Knorr, D., Schröter, K., 2005. Global biomass mapping for an improved understanding of the CO 2 balance - the earth observation mission carbon-3D. Remote Sens. Environ. 94 (1), 94–104. https://doi.org/10.1016/j.rse.2004.09.006.

Houghton, R.A., Hall, F., Goetz, S.J., 2009. Importance of biomass in the global carbon cycle. J. Geophys. Res. Biogeosci. 114 (3), 1–13. https://doi.org/10.1029/2009JG000935.

Hudak, A.T., Evans, J.S., Smith, A.M.S., 2009. LiDAR utility for natural resource managers. Remote Sens. 1 (4), 934–951. https://doi.org/10.3390/rs1040934.

Hummel, S., Hudak, A.T., Uebler, E.H., Falkowski, M.J., Megown, K.A., 2011. A comparison of accuracy and cost of LiDAR versus stand exam data for landscape management on the Malheur National Forest. J. For. 109 (August), 267–273. Retrieved from.

Krieger, G., Hajnsek, I., Papathanassiou, K., Eineder, M., Younis, M., De Zan, F., Lopez-Dekker, P., Huber, S., Werner, M., Prats, P., Fiedler, H., Freeman, A., Rosen, P., Hensley, S., Johnson, W., Veilleux, L., Grafmueller, B., Werninghaus, R., Bamler, R., Moreira, A., 2010. Tandem-L: a mission for monitoring earth system dynamics with high resolution SAR interferometry. In: Proceedings of the European Conference on Synthetic Aperture Radar. EUSAR, pp. 506–509. https://doi.org/10.1109/RADAR.2009.4977077.

Lefsky, M.A., 2010. A global forest canopy height map from the moderate resolution imaging spectroradiometer and the Geoscience Laser Altimeter System. Geophys. Res. Lett. 37, L15401 https://doi.org/10.1029/2010GL043622.

Li, W., Niu, Z., Shang, R., Qin, Y., Wang, L., Chen, H., 2020. High-resolution mapping of forest canopy height using machine learning by coupling ICESat-2 LiDAR with Sentinel-1, Sentinel-2 and Landsat-8 data. Int. J. Appl. Earth Obs. Geoinf. 92 (February), 102163. https://doi.org/10.1016/j.jag.2020.102163.

Liaw, A., Wiener, M., 2015. RandomForest: Breiman and Cutler's Random Forests for Classification and Regression, Version 4.6–12. Available online: https://cran.rproject.org/web/packages/randomForest/.

Lucas, R., Armston, J., Fairfax, R., et al., 2010. An Evaluation of the ALOS PALSAR L-Band Backscatter — Above Ground Biomass Relationship Queensland, Australia: Impacts of Surface Moisture Condition and Vegetation Structure, 3(4), pp. 576–593. https://doi.org/10.1109/JSTARS.2010.2086436.

Mareya, H.T., Tagwireyi, P., Ndaimani, H., Gara, T.W., Gwenzi, D., 2018. Estimating Tree Crown Area and Aboveground Biomass in Miombo Woodlands From, 11(3), pp. 868–875. https://doi.org/10.1109/JSTARS.2018.2799386.

Marshak, C., Simard, M., Denbina, M., 2019. Monitoring forest loss in ALOS/PALSAR time-series with superpixels. Remote Sens. 11 (5), 556. https://doi.org/10.3390/rs11050556.

Mitchard, E.T., Saatchi, S.S., Woodhouse, I.H., Nangendo, G., Ribeiro, N., Williams, M., Ryan, C.M., Lewis, S.L., Feldpausch, T., Meir, P., 2009. Using satellite radar backscatter to predict above-ground woody biomass: a consistent relationship across four different African landscapes. Geophys. Res. Lett. 36, 23. https://doi.org/10.1029/2009GL040692.

Montesano, P.M., Rosette, J., Sun, G., North, P., Nelson, R.F., Dubayah, R.O., Kharuk, V., 2015. Remote sensing of environment the uncertainty of biomass estimates from modeled ICESat-2 returns across a boreal forest gradient. Remote Sens. Environ. 158, 95–109. https://doi.org/10.1016/j.rse.2014.10.029.

Moreira, A., Bachmann, M., Balzer, W., Tridon, D.B., Diedrich, E., Fritz, T., Grigorov, C., Kahle, R., Krieger, G., Hajnsek, I., Huber, S., Jörg, H., Klenk, P., Lachaise, M., Maier, M., Maurer, E., Papathanassiou, K., Parizzi, A., Prats, P., Zonno, M., 2018. Tandem-L: project status and main findings of the phase B1 study. In: International Geoscience and Remote Sensing Symposium (IGARSS), 2018-July(February), pp. 8667–8670. https://doi.org/10.1109/IGARSS.2018.8518591.

Narine, L.L., Popescu, S.C., Malambo, L., 2019a. Synergy of ICESat-2 and Landsat for mapping Forest aboveground biomass with deep learning. Remote Sens. 11, 1503. https://doi.org/10.3390/rs11121503.

Narine, L.L., Popescu, S., Neuenschwander, A., Zhou, T., Srinivasan, S., Harbeck, K., 2019b. Estimating aboveground biomass and forest canopy cover with simulated ICESat-2 data. Remote Sens. Environ. 224, 1–11.

Neuenschwander, A.L., Magruder, L.A., 2019. Canopy and terrain height retrievals with ICESat-2: a first look. Remote Sens. 11 (14), 1721. https://doi.org/10.3390/rs11141721.

Neuenschwander, A., Pitts, K., 2019. Remote sensing of environment the ATL08 land and vegetation product for the ICESat-2 Mission. Remote Sens. Environ. 221 (April 2018), 247–259. https://doi.org/10.1016/j.rse.2018.11.005.

Neumann, M., Saatchi, S., Ulander, L.M.H., Fransson, J.E.S., Mar 2012. Assessing performance of L- and P-band polarimetric interferometric SAR data in estimating boreal forest above-ground biomass. IEEE Trans. Geosci. Remote Sens. 50 (3), 714–726.

NISAR, 2020. NASA-ISRO SAR Mission (NISAR). Disponible at. https://nisar.jpl.nasa.gov/.

Persson, H., Fransson, J.E., 2014. Forest variable estimation using radargrammetric processing of TerraSAR-X images in boreal forests. Remote Sens. 6, 2084–2107.

Qi, W., Lee, S.K., Hancock, S., Luthcke, S., Tang, H., Armston, J., Dubayah, R., 2019a. Improved forest height estimation by fusion of simulated GEDI Lidar data and TanDEM-X InSAR data. Remote Sens. Environ. 221, 621–634. https://doi.org/10.1016/j.rse.2018.11.035.

Qi, W., Saarela, S., Armston, J., Ståhl, G., Dubayah, R., 2019b. Remote sensing of environment forest biomass estimation over three distinct forest types using TanDEM-X InSAR data and simulated GEDI lidar data. Remote Sens. Environ. 232 (August 2018), 111283. https://doi.org/10.1016/j.rse.2019.111283.

Quegan, S., Le Toan, T., Chave, J., Dall, J., Exbrayat, J.F., Minh, D.H.T., Lomas, M., D'Alessandro, M.M., Paillou, P., Papathanassiou, K., Rocca, F., Saatchi, S., Scipal, K., Shugart, H., Smallman, T.L., Soja, M.J., Tebaldini, S., Ulander, L., Villard, L., Williams, M., 2019. The European Space Agency BIOMASS mission: measuring forest above-ground biomass from space. Remote Sens. Environ. 227 (September 2018), 44–60. https://doi.org/10.1016/j.rse.2019.03.032.

Saarela, S., Holm, S., Healey, S.P., Andersen, H.-E., Petersson, H., Prentius, W., Patterson, P.L., Næsset, E., Gregoire, T.G., Ståhl, G., 2018. Generalized hierarchical model-based estimation for aboveground biomass assessment using GEDI and landsat data. Remote Sens. 10 (11), 1832. https://doi.org/10.3390/rs10111832.

Saatchi, S., Marlier, M., Chazdon, R.L., Clark, D.B., Russell, A.E., 2011. Remote sensing of environment impact of spatial variability of tropical forest structure on radar estimation of aboveground biomass. Remote Sens. Environ. 115 (11), 2836–2849. https://doi.org/10.1016/j.rse.2010.07.015.

Santoro, M., Cartus, O., 2018. Research pathways of feorest above-ground biomass estimation based on SAR backscatter and interferometric SAR observations. Remote Sens. 10, 608. https://doi.org/10.3390/rs10040608.

Schlund, M., Scipal, K., Quegan, S., 2018. Assessment of a power law relationship between P-band SAR backscatter and aboveground biomass and its implications for BIOMASS mission performance. In: IEEE Jnl. Selected Topics in Applied Earth Observations and Remote Sensing, 11, pp. 3538–3547. https://doi.org/10.1109/JSTARS.2018.2866868.

Schmitt, M., Zhu, X.X., 2016. Data fusion and remote sensing: an ever-growing relationship. IEEE Geos. Rem. Sens. Mag. 4 (4), 6–23. https://doi.org/10.1109/MGRS.2016.2561021.

Shao, Z., Zhang, L., 2016. Estimating forest aboveground biomass by combining optical and SAR data: a case study in Genhe, Inner Mongolia, China. Sensors 16, 834.

Shimada, M., Member, S., Ohtaki, T., 2010. Generating Large-Scale High-Quality SAR Mosaic Datasets: Application to PALSAR Data for Global Monitoring, 3(4), pp. 637–656.

Silva, C.A., Saatchi, S., Garcia, M., Labri, N., Klauberg, C., Meyer, V., Hudak, A.T., 2018. Comparison of Small- and Large-Footprint Lidar Characterization of Tropical Forest Aboveground Structure and Biomass: A Case Study From Central Gabon, 11(10), pp. 3512–3526.

Silva, C.A., Klauberg, C., Hudak, A.T., Vierling, L.A., Liesenberg, V., Carvalho, S.P.E., Rodriguez, L.C., 2016. A Principal Component Approach for Predicting the Stem Volume in Eucalyptus Plantations in Brazil Using Airborne LiDAR Data, pp. 1–12. https://doi.org/10.1093/forestry/cpw016.

Silveira, E.M.O., Henrique, S., Silva, G., Acerbi-junior, F.W., Carvalho, M.C., Marcelo, L., Wulder, M.A., 2019. Int J Appl Earth Obs Geoinformation Object-based random forest modelling of aboveground forest biomass outperforms a pixel-based approach in a heterogeneous and mountain tropical environment. Int J Appl Earth Obs Geoinformation 78 (December 2018), 175–188. https://doi.org/10.1016/j.jag.2019.02.004.

Simard, M., Riel, B.V., Denbina, M., Hensley, S., 2016. Radiometric correction of airborne radar images over forested terrain with topography. IEEE Trans. Geosci. Remote Sens. 54, 4488–4500. https://doi.org/10.1109/TGRS.2016.2543142.

Sinha, S., Jeganathan, C., Sharma, L.K., Nathawat, M.S., 2015. A Review of Radar Remote Sensing for Biomass Estimation, pp. 1779–1792. https://doi.org/10.1007/s13762-015-0750-0.

Sonoma County Climatic Zones, 2007. University of California Cooperative Extension. Access on August 15, 2019. Available at <https://web.archive.org/web/20070812154328/ http://cesonoma.ucdavis.edu/hortic/sonoma_county_climate.pdf>.

Sun, G., Ranson, K.J., Guo, Z., Zhang, Z., Montesano, P., Kimes, D., 2011. Remote sensing of environment forest biomass mapping from lidar and radar synergies. Remote Sens. Environ. 115 (11), 2906–2916. https://doi.org/10.1016/j.rse.2011.03.021.

Swatantran, A., Tang, H., Barrett, T., Decola, P., Dubayah, R., 2016. Rapid, High-Resolution Forest Structure and Terrain Mapping over Large Areas using Single Photon Lidar, May. Nature Publishing Group, pp. 1–12. https://doi.org/10.1038/srep28277.

Tian, X., Su, Z., Chen, E., Li, Z., van der Tol, C., Guo, J., He, Q., 2012. Estimation of forest above-ground biomass using multi-parameter remote sensing data over a cold and arid area. Int. J. Appl. Earth Obs. Geoinf. 14 (1), 160–168. https://doi.org/10.1016/j.jag.2011.09.010.

Wilcoxon, F., 1945. Individual comparisons by ranking methods. Biom. Bull. 1, 80–83.

Yu, Y., Saatchi, S., 2016. Sensitivity of L-band SAR backscatter to aboveground biomass of global forests. Remote Sens. 8 (6), 522. https://doi.org/10.3390/rs8060522.

Zhang, C., Denka, S., Cooper, H., Mishra, D.R., 2018. Remote sensing of environment quantification of sawgrass marsh aboveground biomass in the coastal Everglades using object-based ensemble analysis and Landsat data. Remote Sens. Environ. 204 (September 2017), 366–379. https://doi.org/10.1016/j.rse.2017.10.018.

Zolkos, S.G., Goetz, S.J., Dubayah, R., 2013. Remote sensing of environment A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. Remote Sens. Environ. 128, 289–298. https://doi.org/10.1016/j.rse.2012.10.017.